Identification of similar seismic events using a phase-only correlation technique

著者	Hirokazu Moriya
journal or publication title	Construction and Building Materials
volume	168
page range	984-987
year	2019-03-12
URL	http://hdl.handle.net/10097/00127713

doi: 10.1016/j.conbuildmat.2018.03.053

Identification of similar seismic events using a phase-only correlation technique

Hirokazu Moriya

Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan

ABSTRACT: This paper proposes a method of seismic cluster analysis using phase-only correlation (POC) to both identify and hierarchically classify similar acoustic emission (AE) waveforms. The POC value of a time-varying spectral representation is used to evaluate the similarity between two images of a waveform, and cluster analysis is used to classify waveforms into groups according to a distance measure. This method is applied to waveforms from local earthquakes in Japan and assessed by its ability to identify similar waveforms perturbed by white noise.

1 INTRODUCTION

Acoustic emission (AE) events with similar waveforms are considered to be repeating events on the same crack surface. Precise source location techniques that use cross-spectrum analysis and cross-correlation methods are applicable to similar AE events, making it possible to identify the orientation of crack planes as well as their locations [1]. Cross-correlation functions among AE events are often used to identify similar AE waveforms. However, it is difficult to analyze and classify noisy waveforms using the degree of similarity.

Phase-only correlation (POC) is used in electrical and communication engineering for image matching and is a simple and robust technique for evaluating the similarity of two images [2, 3, 4]. The POC method deals well with noise and perturbations superimposed on signals [4] and has great potential for use in geophysical exploration, for instance, for identification of similar seismic events, similar seismic traces, and differences in seismic images. Here I describe the application of a technique based on 2-dimensional (2D) POC to evaluate and classify the waveform similarity of AE events.

2 METHOD

The POC method was introduced in the field of pattern recognition of images such as fingerprints. The POC function F_h is defined by the following equations [4]:

$$\begin{aligned} F_h(k_1,k_2) &= \sum_{n_1=-M_1}^{M_1} \sum_{n_2=-M_2}^{M_2} f_h(n_1,n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_h(k_1,k_2) e^{j\theta_h(k_1,k_2)} \end{aligned}, \quad (1)$$

where $k_1 = -M_1,...,M_1$, $k_2 = -M_2,...,M_2$, $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$, and $W_{N_2} = e^{-j\frac{2\pi}{N_2}}$. $A_h(k_1,k_2)$ denotes the amplitudes of $F_h(k_1,k_2)$, and $e^{j\theta_h(k_1,k_2)}$ is their phase. Then, the function $\hat{R}_{h\ell}(k_1,k_2)$ can be defined as

$$\hat{R}_{h\ell}(k_1,k_2) = \frac{F_h(k_1,k_2)\overline{F}_\ell(k_1,k_2)}{\left|F_h(k_1,k_2)\overline{F}_\ell(k_1,k_2)\right|} = e^{j\theta(k_1,k_2)}, \quad (2)$$

where $\overline{F}_\ell(k_1,k_2)$ represents the complex conjugate of $F_h(k_1,k_2)$ and $\theta(k_1,k_2) = \theta_h(k_1,k_2) - \theta_\ell(k_1,k_2)$. The POC function can be represented by the inverse discrete Fourier transform as follows:

$$\hat{r}_{h\ell}(n_1,n_2)=\frac{1}{N_1N_2}\sum_{k_1=-M_1}^{M_1}\sum_{k_2=-M_2}^{M_2}\hat{R}_{h\ell}(k_1,k_2)W_{N_1}^{-k_1n_1}W_{N_2}^{-k_2n_2}\ , \tag{3}$$

where the value of $\hat{r}_{h\ell}(n_1,n_2)$ ranges from 0 to 1. The value of $\hat{r}_{h\ell}(n_1,n_2)$ is a measure of the similarity of two images, and the position of this peak indicates the translational displacement between the two images.

The peak value of the POC function can also be used for seismic cluster analysis. For this purpose I define a matrix consisting of the peak values of the POC function calculated for all combinations of two waveforms. For a total of Q waveforms, the matrix is defined as follows:

$$\mathbf{X}=\left[R_{h\ell}\right] \tag{4}$$

$$R_{h\ell}=Max\,[\hat{r}_{h\ell}(n_1,n_2)] \qquad (h,\ell=1,...,Q)\,. \tag{5}$$

The rows of **X**, that is, $R_{h1}, R_{h2}, ..., R_{h\ell}$, correspond to observation vectors, which are the peak value of the POC function calculated between the *h*-th and *l*-th seismic events. A distance measure is needed for the cluster analysis to sort the observations into subsets. The city-block distance was used for the distance measure, with the distance between the *s*-th and *t*-th row vectors given by

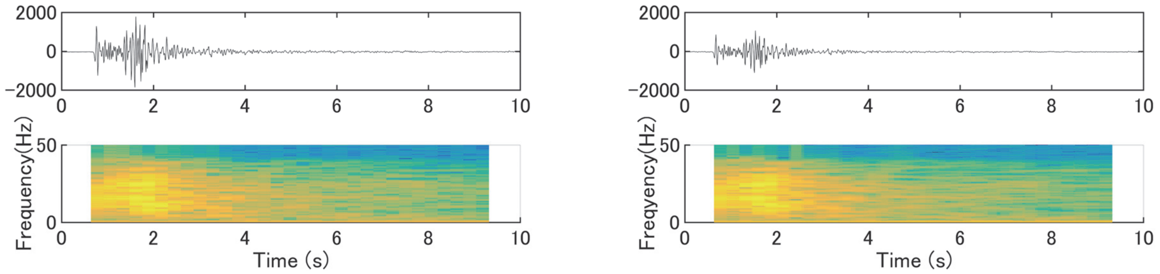
$$d_{st}=\sum_{i=1}^{L}\left|R_{si}-R_{ti}\right|\,. \tag{6}$$

A rule is also needed to determine the distance between clusters when several seismic events have been linked together [5]. For this I used Ward's method, which maximizes at each step the ratio between the variances within clusters and the variances between clusters [6, 7]. This analysis clarifies the linkages among seismic events as well as the grouping of similar seismic events, and the results are presented on a hierarchical tree as distinct branches.

3 APPLICATION TO AE EVENTS

Two-dimensional POC of time-varying spectral representations (TVSPs) of seismic data makes it possible to evaluate in the time–frequency domain the similarity of waveforms for which both the frequency content and signal amplitude change with time. I used the TVSP of a seismic signal to represent a waveform image [8], corresponding to the $f_h(n_1,n_2)$ term in equation (1). Because the TVSP resolves the AE into frequency components that change with time, TVSPs can express the characteristics of an AE wave represented by the multiplication of the source function and the transfer function of the wave propagation path. I calculated the POC function by using TVSPs of seismic events for all the combinations of events. Because the POC function evaluates the similarity between two images and is calculated using the TVSP of waveforms, the evaluation can consider the similarity of frequency content, frequency distribution, and wave attenuation with time, even though the cross-correlation function and the coherency function cannot be used to evaluate changes in waveform characteristics with time, such as decreasing amplitude with time. Figures 1(a) and 1(b) show the waveforms of two local earthquakes, which occurred west of Sendai, Japan at a depth of about 12 km. The TVSP was calculated using the FFT algorithm, and the time window for calculation was moved along the time axis. Figure 2 shows the POC function comparing the two events in Figure 1, and Figure 3 shows the 2D cross-correlation of the two events using the time-varying spectra. The 1D and 2D cross-correlation functions are often used to evaluate the similarity of time series data and two-dimensional images, respectively. The POC function has a sharper and narrower peak than the cross-correlation function, making it easier to judge the similarity of two waveforms [8]. One advantage of the POC function is that it can identify similar images perturbed by noise, and it is more robust because the POC evaluates the 3D shape of time-varying spectra. In contrast, the 2D cross-correlation function in Figure 3 has a broad envelope and displays periodic peaks along the frequency (n_1) axis, the result of multiplication by a time window.

Because the POC method uses time-varying spectra of waveforms, the waveforms are decomposed into the time and frequency domains. Therefore, the similarity of waveforms is evaluated in both time and frequency domains. In my studies using simulated waveforms, the POC had a maximum value larger than 0.6 at a signal-to-noise ratio of –20 dB, whereas 1D and 2D cross-correlation functions had maximum values of 0.4.



(a) (b)

Figure 1: Two waveforms and their time-varying spectra, analyzed in Figures 2 and 3.

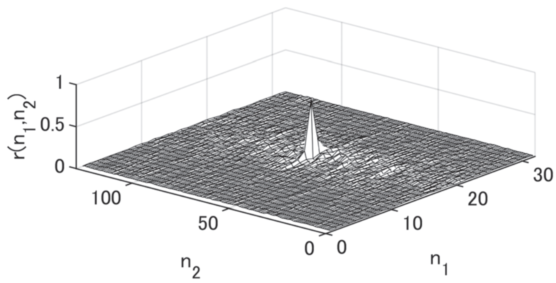


Figure 2: POC of waveforms using time-varying spectra.

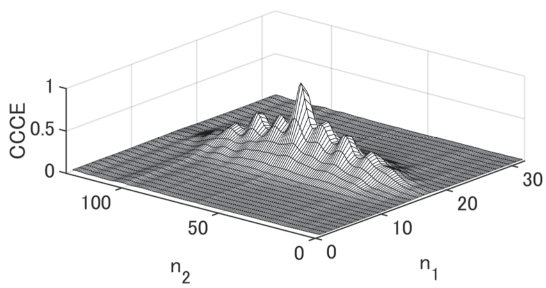


Figure 3: 2D cross-correlation of waveforms.

The POC method was applied to a group of local earthquakes. Figure 4 shows a set of waveforms used for classification, of which the first five were manually judged to be a group of similar events. The POC function was calculated between all pairs of these events and subjected to cluster analysis, using the peak values of POC functions to evaluate the distance of two events. Figure 5 shows the result of the cluster analysis. The pairs or groups of waveforms that are identified as similar are connected with a horizontal line, the position of which on the vertical axis represents their similarity.

This example suggests that cluster analysis can identify similar waveforms and aggregate them using the distance measure defined by the peak value of the POC function. Conventional analysis has no appropriate method to link and classify pairs and groups of similar events, even though the similarity of pairs of waveforms can be evaluated using cross-correlation or coherence functions. The method presented here can both identify and group similar seismic events, and it can be automated because few parameters are needed for the

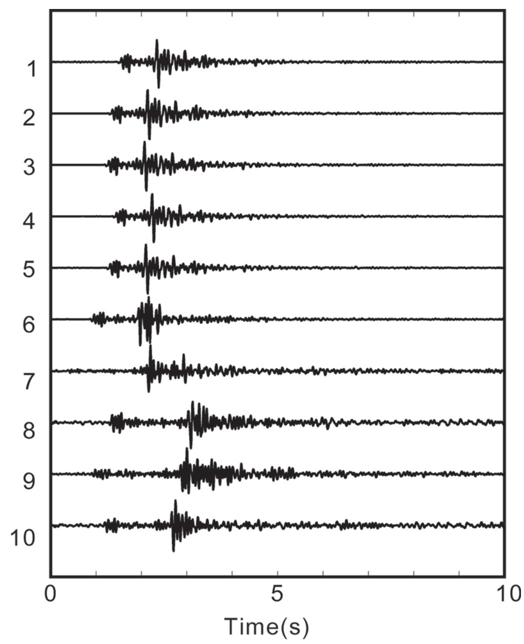


Figure 4: Example of waveforms.

calculations. This method can be also used for sequential analysis to find similar seismic events and assign them one by one into groups.

4 CONCLUSION

This paper presents a method that uses the POC function to evaluate the similarity of seismic events and uses the peak value of POC as a distance measure for hierarchical cluster analysis. Applications to seismic waveforms show that the method can detect similar events and classify them into branches of a tree indicating their degree of linkage, and that it can distinguish waveforms for sorting into different groups.

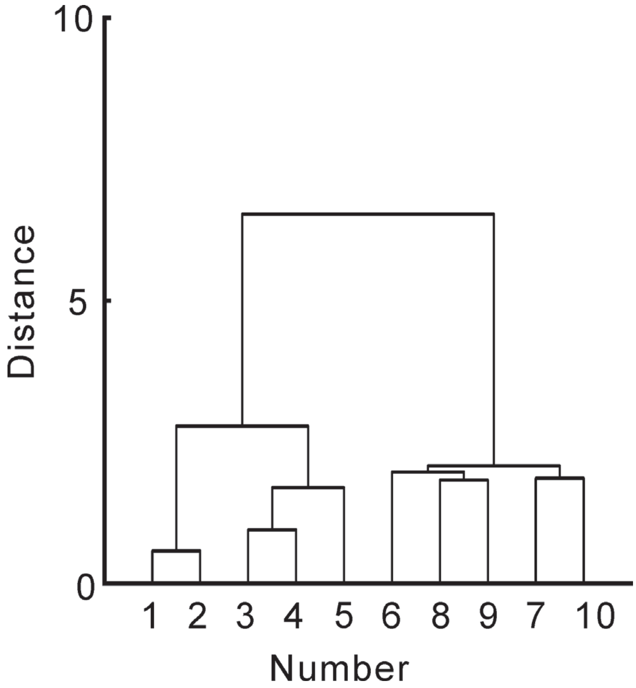


Figure 5: Result of classification.

ACKNOWLEDGEMENTS

Part of this work was supported by Grants-in-Aid for Scientific Research (A) (26249137) from the Ministry of Education, Science, Sports and Culture. I thank the National Research Institute for Earth Science and Disaster Prevention (NIED) of Japan for providing the earthquake waveform data, which were downloaded from the NIED website (http://www.hinet.bosai.go.jp/).

REFERENCES

[1] Moriya, H., Niitsuma, H., Baria, R., (2003). “Multiplet-clustering analyses reveals structural details within the seismic cloud at the Soultz geothermal field, France”, *Bulletin of the Seismological Society of America*, Vol.93, pp.1606–1620.
[2] Maltoni, D., Maio, D., Jain, A. K., Prabhakar, S., (2003). “Handbook of fingerprint recognition”, Springer.
[3] Takita, K., Aoki, T., Sasaki, Y., Higuchi, T., Kobayashi, K., (2003). “High-accuracy subpixel image registration based on phase-only correlation”, *IEICE Transaction Fundamentals*, Vol.E86-A, pp.1925–1934.
[4] Ito, K., Nakajima, H., Kobayashi, K., Aoki, T., Higuchi, T., (2004). “A fingerprint matching algorithm using phase-only correlation”, *IEICE Transactions. Fundamentals,* Vol. E87-A, pp.682–691.
[5] Weatherill, G., Burton, P.W., (2009). “Delineation of shallow seismic source zones using K-means cluster analysis, with application to the Aegean region”, *Geophysical Journal International*, Vol. 176, pp.565-588.
[6] Ward, J.H., (1963). “Hierarchical grouping to optimize an objective function”, *Journal of the American Statistical Association*, Vol.58, pp.236–244.
[7] Tan, P-G, Steinbach, M., Kumar V., (2006). “Introduction to data mining”, *Addison-Wesley*.
[8] Moriya, H., (2011). “Phase-only correlation for identification of similar seismic events”, *Geophysics*, pp.WC37-WC45.